Recent Developments in Matching with Constraints

By Yuichiro Kamada and Fuhito Kojima*

The medical matching market in Japan is subject to *constraints*: there is a constraint on the number of doctors that can be matched to each prefecture (partitioning the country into 47). This feature differentiates itself from the standard two-sided matching à la Gale and Shapley (1962) because these "regional cap" constraints can be violated even if all hospital capacities are respected. As a consequence, there may not exist a stable matching in the standard sense which ignores the regional caps.

In Kamada and Kojima (2015), we introduced this problem, proposed and analyzed a possible solution, and discussed other various real-market applications of our result.[1] Such examples include Chinese graduate-school admissions, UK medical match, and Scotland's matching between teachers and schools. In subsequent papers, we continued to work on matching with various forms of constraints and obtained some general insights. This paper aims to reorganize some of our main findings from this effort, based on Kamada and Kojima (2016*a*,*b*).[2]

We start our discussion by considering a simple constraint structure, and then move on to a more complex case. In Section II, we consider the situation in which the only information available to the market designer is about hard constraints. That is, upper-bound constraints are imposed on certain subsets of hospitals. In that model, we analyze properties of two stability concepts that take constraints into account.

The set of agents on which a constraint is imposed is often associated with a certain stakeholder. For example, in the Japanese medical match, constraints are imposed on prefectures, and each prefecture has its own association of doctors. In Section III, we consider the situation where the designer has information about each region's preferences—called "regional preferences"—over the allocations of doctors among the hospitals in it.

I. Preliminary Definitions

Let there be a finite set of doctors D and a finite set of hospitals H. Each doctor d has a strict preference relation $\succ_d$ over $H \cup \{\emptyset\}$, where $\emptyset$ denotes the outside option. Each hospital h has a strict preference relation $\succ_h$ over 2^D. We denote by $\succ = (\succ_i)_{i \in D \cup H}$ the preference profile of all doctors and hospitals.

Each hospital $h \in H$ is endowed with a (physical) **capacity** q_h, which is a nonnegative integer. We assume that preferences of each hospital h are responsive with capacity q_h (Roth, 1985) throughout the paper.

A **matching** μ is a mapping that satisfies (i) $\mu_d \in H \cup \{\emptyset\}$ for all $d \in D$, (ii) $\mu_h \subseteq D$ for all $h \in H$, and (iii) for any $d \in D$ and $h \in H$, $\mu_d = h$ if and only if $d \in \mu_h$. That is, a matching simply specifies which doctor is assigned to which hospital (if any).

A matching μ is **individually rational** if (i) for each $d \in D$, $\mu_d \succ_d \emptyset$ or $\mu_d = \emptyset$, and (ii) for each $h \in H$, $d \succ_h \emptyset$ for all $d \in \mu_h$, and $|\mu_h| \leq q_h$.[3] That is, no agent is matched with an unacceptable partner and each hospital's capacity is respected.

Given matching μ, a pair (d, h) of a doctor and a hospital is called a **blocking pair** if $h \succ_d \mu_d$ and either (i) $|\mu_h| < q_h$ and $d \succ_h \emptyset$, or (ii) $d \succ_h d'$ for some $d' \in \mu_h$. In words, a blocking pair is a pair of a doctor and a hospital who want to be matched with each other

* Kamada: Haas School of Business, University of California, Berkeley, Berkeley, CA 94720, y.cam.24@gmail.com. Kojima: Department of Economics, Stanford University, Stanford, CA 94305, fkojima@stanford.edu. Akhil Vohra and Qihang Wu provided excellent research assistance.

[1] In the present paper, "constraints" refer to upper-bound constraints. Section IV discusses lower-bound constraints.

[2] Much of the notation and texts are borrowed from those papers.

[3] We denote singleton set $\{x\}$ by x when there is no confusion.

(possibly rejecting their partners in the prescribed matching) rather than following the proposed matching.

II. Model without Regional Preferences

Regions. A collection $R \subseteq 2^H \setminus \{\emptyset\}$ is called a set of **regions**. Assume $H \in R$ and $\{h\} \in R$ for all $h \in H$. Each region $r \in R$ is endowed with a nonnegative integer κ_r called a **regional cap**. We denote by $\kappa = (\kappa_r)_{r \in R}$ the profile of regional caps across all regions in R. A matching is **feasible** if $|\mu_r| \leq \kappa_r$ for all $r \in R$, where $\mu_r = \cup_{h \in r} \mu_h$. In other words, feasibility requires that the regional cap for every region is satisfied.

Strong and Weak Stability. The standard definition of stability without regional caps requires individual rationality and the absence of blocking pairs. With regional caps, however, there are cases in which every feasible and individually rational matching admits a blocking pair. For this reason, we allow for the presence of some blocking pairs. To keep the spirit of stability, however, we require certain kinds of blocking pairs not to exist.

A pair of a doctor d and a hospital h is infeasible if moving d to h while keeping other parts of the matching unchanged leads to a violation of a regional cap. To the extent that regional caps encode what matchings are allowed in a given situation, blocking pairs violating a regional cap do not have as much normative support as others. For this reason, our stability concept allows for infeasible blocking pairs to remain.

Formally, given a matching μ, denote by $\mu^{d \to h}$ the matching such that $\mu^{d \to h}_{d'} = \mu_{d'}$ for all $d' \in D \setminus \{d\}$ and $\mu^{d \to h}_d = h$. We say that a pair (d, h) is **infeasible** at μ if $\mu^{d \to h}$ is not feasible.

Definition 1. A matching μ is **strongly stable** if it is feasible, individually rational, and if (d, h) is a blocking pair then $d' \succ_h d$ for all doctors $d' \in \mu_h$ and (d, h) is infeasible at μ.

We say that (d, h) is **weakly infeasible** at μ if $\tilde{\mu}$ is infeasible where $\tilde{\mu}$ is a matching such that $\tilde{\mu}_d = \mu_d$ for all $d \in D$ and one more (hypothetical) doctor is added to h.

Definition 2. A matching μ is **weakly stable** if it is feasible, individually rational, and if (d, h) is a blocking pair then $d' \succ_h d$ for all doctors $d' \in \mu_h$ and (d, h) is weakly infeasible at μ.

Results. The first result pertains to the existence. Although strong stability may be a natural desideratum for a practitioner, the following theorem limits the use of strong stability, while leaving a hope for weak stability.

Theorem 1 (Kamada and Kojima (2016*b*)).

(1) *There does not necessarily exist a strongly stable matching.*

(2) *There exists a weakly stable matching.*

Kamada and Kojima (2016*b*) prove a stronger claim than part (1) of the above theorem. Specifically, a strongly stable matching is guaranteed to exist if and only if the regional constraint is trivial, meaning that any region (with a strictly positive cap) contains only one hospital.

The next theorem provides a justification for weak stability.

Theorem 2 (Kamada and Kojima (2016*b*)).

(1) *A weakly stable matching is constrained efficient.*[4]

(2) *A matching is weakly stable if and only if it is individually rational, feasible, nonwasteful, and satisfies the no-justified-envy property.*[5]

III. Model with Regional Preferences

We now consider the situation in which the market designer is subject not only to regional caps, but also to a governmental goal regarding allocation of doctors. We use such a governmental goal to define a stability concept that is weaker than strong stability so that existence still holds under certain conditions, but stronger than weak stability which implies that all the desirable properties

[4]Constrained efficiency means that there is no feasible matching μ' such that $\mu'_i \succeq_i \mu_i$ for all $i \in D \cup H$ and $\mu'_i \succ_i \mu_i$ for some $i \in D \cup H$.

[5]The last two conditions are defined in Kamada and Kojima (2016*b*).

stated in part (2) of Theorem 2 still hold.

Regional Structure. Let $S \subset R$ be a partition of r. We call S a **largest partition of** r if there exists no other partition $S' \subset R$ of r such that $r' \in S$ implies $r' \subseteq r''$ for some $r'' \in S'$. Let $\mathcal{LP}(r)$ denote the collection of largest partitions of r (note that there can be more than one largest partition of a given region r). For $r \in R$ and $S \in \mathcal{LP}(r)$, we refer to each element of S as a **subregion** of r with respect to S. A set of regions R is a **hierarchy** if $r, r' \in R$ implies $r \subseteq r'$ or $r' \subseteq r$ or $r \cap r' = \emptyset$.

Regional Preferences. When a given region is faced with applications by more doctors than the regional cap, the region has to allocate limited seats among its subregions. We consider the situation in which regions have policy goals in terms of doctor allocations, and formalize such policy goals using the concept of "regional preferences."

For each $r \in R$ that is not a singleton set and $S \in \mathcal{LP}(r)$, a regional preference for r, denoted $\trianglerighteq_{r,S}$, is a weak ordering over $W_{r,S} := \{w = (w_{r'})_{r' \in S} | w_{r'} \in \mathbb{Z}_+ \text{ for every } r' \in S\}$. Vectors such as w are interpreted to be supplies of acceptable doctors to the subregions of region r, but they only specify how many acceptable doctors apply to hospitals in each subregion and provide no information as to who these doctors are. We denote by $\trianglerighteq$ the profile $(\trianglerighteq_{r,S})_{r \in R, S \in \mathcal{LP}(r)}$. We assume that the regional preferences $\trianglerighteq_{r,S}$ satisfy $w \trianglerighteq_{r,S} w'$ and $w' \not\trianglerighteq_{r,S} w$ if $w' \lneq w$. This condition formalizes the idea that region r prefers to fill as many positions in its subregions as possible.

Given $\trianglerighteq_{r,S}$, a function

$$\tilde{\mathrm{Ch}}_{r,S} : W_{r,S} \times \mathbb{Z}_+ \to W_{r,S}$$

is an **associated quasi choice rule** if $\tilde{\mathrm{Ch}}_{r,S}(w;t) \in \arg\max_{\trianglerighteq_{r,S}} \{w' | w' \leq w, \sum_{r' \in S} w'_{r'} \leq t\}$ for any non-negative integer vector $w = (w_{r'})_{r' \in S}$ and non-negative integer t. Intuitively, $\tilde{\mathrm{Ch}}_{r,S}(w,t)$ is a best vector of numbers of doctors allocated to subregions of r given a vector of numbers w under the constraint that the sum of the numbers of doctors cannot exceed the quota t.

We say that $\trianglerighteq_{r,S}$ is **substitutable** if there exists an associated quasi choice rule $\tilde{\mathrm{Ch}}_{r,S}$ that satisfies

$$w \leq w', t \geq t' \Rightarrow \tilde{\mathrm{Ch}}_{r,S}(w;t) \geq \tilde{\mathrm{Ch}}_{r,S}(w';t') \wedge w.$$

Throughout our analysis, we assume that $\trianglerighteq_{r,S}$ is substitutable for any $r \in R$ and $S \in \mathcal{LP}(r)$.

Stability. For $R' \subseteq R$, we say that μ is **Pareto superior** to μ' for R' if $(|\mu_{r'}|)_{r' \in S} \trianglerighteq_{r,S} (|\mu'_{r'}|)_{r' \in S}$ for all (r, S) where $r \in R'$ and $S \in \mathcal{LP}(r)$, with at least one of the relations holding strictly.

We say that a pair (d,h) is **illegitimate** at μ if there exists $r \in R$ with $|\mu_r| = \kappa_r$ such that $\mu^{d \to h}$ is not Pareto superior to μ for $\{r' \in R \mid \mu_d, h \in r' \text{ and } r' \subseteq r\}$.

Definition 3. A matching μ is **stable** if it is feasible, individually rational, and if (d,h) is a blocking pair then $d' \succ_h d$ for all doctors $d' \in \mu_h$ and (d,h) is either infeasible or illegitimate at μ.

A doctor-hospital pair is illegitimate if the movement of doctor d to h does not lead to a Pareto superior distribution of doctors for a certain set of regions.We require any region r' in this set to satisfy two conditions. First, r' has to contain both hospitals μ_d (the original hospital for d) and h, as this corresponds to the case in which r' has a stake in the distributions of doctors involving these hospitals. Second, the region r' should be currently "constrained." That is, it is a subset of some region r whose regional cap is full in the present matching: It is in such a case that the region r should ration the distribution of doctors among its subregions, each of which needs to ration the distribution among its subregions, and so forth, thus indirectly constraining the number of doctors that can be matched in r'.

Practically, moving a doctor from one hospital to another involves administrative tasks on the part of relevant regions, hence disallowing only those blocking pairs that Pareto-improve the relevant regions is, in our view, the most plausible notion in our environment.[6]

[6] An alternative notion of illegitimacy may be to regard a doctor-hospital pair as illegitimate if moving them leads to a Pareto inferior distribution of doctors for the set of regions that we consider here. As detailed in Kamada and Kojima

It is straightforward from the definition of stability that strong stability implies stability, and stability implies weak stability. Kamada and Kojima (2016*b*) further show the following.

Proposition 1 (Kamada and Kojima (2016*b*))**.**

(1) *A matching is strongly stable if and only if it is stable for all possible regional preferences.*

(2) *A matching is weakly stable if and only if it is stable under some regional preference profile.*

Mechanism. Fixing $(\kappa, \trianglerighteq)$, a mechanism is defined as a mapping from preference profits to matchings. Stability and strategy-proofness for doctors of a mechanism are defined in the standard manner.

Theorem 3 (Kamada and Kojima (2016*a*))**.** *Fix D with $|D| \geq 2$, H, and a set of regions R. The following statements are equivalent.*

(1) *R is a hierarchy.*

(2) *For each $(\kappa, \trianglerighteq)$, there exists a mechanism that is stable and strategy-proof for doctors.*

The theorem identifies the conditions on the markets for which we can find a mechanism that is stable and strategy-proof for doctors. Since our proof for the assertion that statement (1) implies statement (2) is constructive, for those markets in which constraints are a hierarchy, the theorem tells us that we can directly use the mechanism we construct. Also, for those markets in which constraints do not form a hierarchy, the theorem shows that there is no hope of adopting a mechanism that is stable and strategy-proof for doctors.

IV. Discussions

Matching with constraints is a new topic in market design. As such, many aspects of it are waiting to be investigated.

One open question is whether there is a mechanism that is weakly stable and strategy-proof for doctors. Although an affirmative answer follows from Theorem 3 under hierarchical constraints, the answer to this question is unknown for a more general class of constraints.[7]

Another possible direction of research is to study more general feasibility constraints. Kamada and Kojima (2016*b*) allow for a broader class of upper-bound constraints than those presented here (Goto et al. (2016*b*) also study that class of constraints). This class of constraints, however, excludes lower-bound (floor) constraints, which are studied by Ehlers et al. (2014), Goto et al. (2016*a*), and Fragiadakis and Troyan (2016), among others. Investigating possible connections between those papers and ours await future research.

Finally, it will be interesting to study the connection between our model of matching with constraints and other related models. They include matching where the choices of *individual hospitals* are subject to constraints, such as Roth (1991), Abdulkadiroğlu and Sönmez (2003), Hafalir, Yenmez and Yildirim (2013), Westkamp (2013), Sönmez (2013), and Kominers and Sönmez (2016). Biro et al. (2010) consider a constrained matching market with additional assumptions on preferences. Budish et al. (2013) study object allocation under hierarchical constraints. Technically, Kamada and Kojima (2016*a*) provide a method to make a connection between matching with constraints and matching with contracts (Hatfield and Milgrom, 2005; Hatfield and Kominers, 2012), and similar techniques can be employed to reproduce results in some of the papers cited above.[8] However, a general theory that provides a unified understanding of various types of constraints is still open.

REFERENCES

Abdulkadiroğlu, Atila, and Tayfun

(2016*a*), this notion has several problematic features.

[7]Goto et al. (2016*b*) propose a mechanism that is strategy-proof for doctors in the same environment as Kamada and Kojima (2016*b*), but the outcome of their mechanism is not weakly stable.

[8]Kojima, Tamura and Yokoo (2015) explore this connection through the theory of discrete convex analysis.

Sönmez. 2003. "School Choice: A Mechanism Design Approach." *American Economic Review*, 93: 729–747.

Biro, P., T. Fleiner, R.W. Irving, and D.F. Manlove. 2010. "The College Admissions Problem with Lower and Common Quotas." *Theoretical Computer Science*, 411(34-36): 3136–3153.

Budish, Eric, Yeon-Koo Che, Fuhito Kojima, and Paul R. Milgrom. 2013. "Designing Random Allocation Mechanisms: Theory and Applications." *American Economic Review*, 103(2): 585–623.

Ehlers, Lars, Isa E. Hafalir, M. Bumin Yenmez, and Muhammed A. Yildirim. 2014. "School Choice with Controlled Choice Constraints: Hard Bounds versus Soft Bounds." *Journal of Economic Theory*, 153: 648–683.

Fragiadakis, Daniel, and Peter Troyan. 2016. "Improving Matching under Hard Distributional Constraints." forthcoming, *Theoretical Economics.*

Gale, David, and Lloyd S. Shapley. 1962. "College Admissions and the Stability of Marriage." *American Mathematical Monthly*, 69: 9–15.

Goto, Masahiro, Atsushi Iwasaki, Yujiro Kawasaki, Ryoji Kurata, Yosuke Yasuda, and Makoto Yokoo. 2016*a*. "Strategyproof matching with regional minimum and maximum quotas." *Artificial Intelligence*, 235: 40–57.

Goto, Masahiro, Fuhito Kojima, Ryoji Kurata, Akihisa Tamura, and Makoto Yokoo. 2016*b*. "Designing Matching Mechanisms under General Distributional Constraints." *American Economic Journal: Microeconomics, forthcoming.*

Hafalir, Isa E, M Bumin Yenmez, and Muhammed A Yildirim. 2013. "Effective affirmative action in school choice." *Theoretical Economics*, 8(2): 325–363.

Hatfield, John William, and Paul Milgrom. 2005. "Matching with Contracts." *American Economic Review*, 95: 913–935.

Hatfield, John William, and Scott Duke Kominers. 2012. "Matching in Networks with Bilateral Contracts." *American Economic Journal: Microeconomics*, 4(1): 176–208.

Kamada, Yuichiro, and Fuhito Kojima. 2015. "Efficient Matching under Distributional Constraints: Theory and Applications." *American Economic Review*, 105(1): 67–99.

Kamada, Yuichiro, and Fuhito Kojima. 2016*a*. "Stability and Strategy-Proofness for Matching with Constraints: A Necessary and Sufficient Condition." mimeo.

Kamada, Yuichiro, and Fuhito Kojima. 2016*b*. "Stability Concepts in Matching with Distributional Constraints." forthcoming, *Journal of Economic Theory.*

Kojima, Fuhito, Akhisa Tamura, and Makoto Yokoo. 2015. "Designing Matching Mechanisms under Constraints: An Approach from Discrete Convex Analysis." mimeo.

Kominers, Scott Duke, and Tayfun Sönmez. 2016. "Matching with slot-specific priorities: Theory." *Theoretical Economics*, 11(2): 683–710.

Roth, Alvin E. 1985. "The College Admission Problem is not Equivalent to the Marriage Problem." *Journal of Economic Theory*, 36: 277–288.

Roth, Alvin E. 1991. "A Natural Experiment in the Organization of Entry Level Labor Markets: Regional Markets for New Physicians and Surgeons in the U.K." *American Economic Review*, 81: 415–440.

Sönmez, Tayfun. 2013. "Bidding for Army Career Specialties: Improving the ROTC Branching Mechanism." *Journal of Political Economy*, 121(1): 186–219.

Westkamp, Alexander. 2013. "An Analysis of the German University Admissions System." *Economic Theory*, 53(3): 561–589.